

PCT/FR 03/02770 #2
Rec'd PCT/PTO 08 MAR 2005

REC'D 01 DEC 2003

WIPO PCT

BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 30 SEP. 2003**DOCUMENT DE PRIORITÉ****PRÉSENTÉ OU TRANSMIS
CONFORMÉMENT À LA
RÈGLE 17.1.a) OU b)**

Pour le Directeur général de l'Institut
national de la propriété industrielle
Le Chef du Département des brevets

Martine PLANCHE

INSTITUT
NATIONAL DE
LA PROPRIÉTÉ
INDUSTRIELLE

SIEGE
26 bis, rue de Saint Petersburg
75800 PARIS cedex 08
Téléphone : 33 (0)1 53 04 53 04
Télécopie : 33 (0)1 53 04 45 23
www.inpi.fr

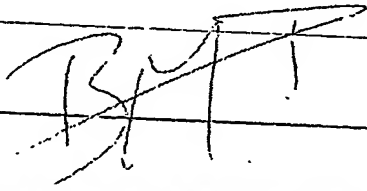
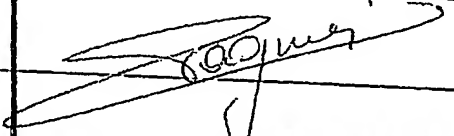
REQUÊTE EN DÉLIVRANCE 1/2

Réservé à l'INPI

Cet imprimé est à remplir lisiblement à l'encre noire

DB 540 W / 260899

REMISE DES PIÈCES DATE 24 SEPT 2002 LIEU 75 INPI PARIS N° D'ENREGISTREMENT 0211789 NATIONAL ATTRIBUÉ PAR L'INPI DATE DE DÉPÔT ATTRIBUÉE 24 SEP. 2002 PAR L'INPI		1 NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE Madame Pascale BROCHARD THALES Intellectual Property 13, avenue du Président Salvador Allende 94117 ARCUEIL Cedex	
Vos références pour ce dossier (facultatif) 62877			
Confirmation d'un dépôt par télécopie <input type="checkbox"/> N° attribué par l'INPI à la télécopie			
2 NATURE DE LA DEMANDE		Cochez l'une des 4 cases suivantes	
Demande de brevet		<input checked="" type="checkbox"/>	
Demande de certificat d'utilité		<input type="checkbox"/>	
Demande divisionnaire		<input type="checkbox"/>	
Demande de brevet initiale N° _____ Date ____/____/____ ou demande de certificat d'utilité initiale N° _____ Date ____/____/____			
Transformation d'une demande de brevet européen Demande de brevet initiale N° _____ Date ____/____/____			
3 TITRE DE L'INVENTION (200 caractères ou espaces maximum) PROCEDE DE RECONNAISSANCE VOCALE AVEC CORRECTION AUTOMATIQUE			
4 DÉCLARATION DE PRIORITÉ OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE FRANÇAISE		Pays ou organisation _____ N° _____ Date ____/____/____ Pays ou organisation _____ N° _____ Date ____/____/____ Pays ou organisation _____ N° _____ Date ____/____/____ <input type="checkbox"/> S'il y a d'autres priorités, cochez la case et utilisez l'imprimé «Suite»	
5 DEMANDEUR		<input type="checkbox"/> S'il y a d'autres demandeurs, cochez la case et utilisez l'imprimé «Suite»	
Nom ou dénomination sociale		THALES	
Prénoms			
Forme juridique		Société Anonyme	
N° SIREN		5 5 2 0 5 9 0 2 4 :	
Code APE-NAF			
Adresse	Rue	173, boulevard Haussmann	
	Code postal et ville	75008	PARIS
Pays		FRANCE	
Nationalité		Française	
N° de téléphone (facultatif)			
N° de télécopie (facultatif)			
Adresse électronique (facultatif)			

REMISE DES PIÈCES DATE 24 SEPT 2002 LIEU 75 INPI PARIS N° D'ENREGISTREMENT 0211789 NATIONAL ATTRIBUÉ PAR L'INPI		Réservé à l'INPI		DB 540 W / 2508
Vos références pour ce dossier : (facultatif)		62877		
6 MANDATAIRE				
Nom		BROCHARD		
Prénom		Pascale		
Cabinet ou Société		THALES		
N° de pouvoir permanent et/ou de lien contractuel		8325		
Adresse	Rue	13, avenue du Président Salvador Allende		
	Code postal et ville	94117	ARCUEIL Cedex	
N° de téléphone (facultatif)		01 41 48 45 67		
N° de télécopie (facultatif)		01 41 48 45 01		
Adresse électronique (facultatif)				
7 INVENTEUR (S)				
Les inventeurs sont les demandeurs		<input type="checkbox"/> Oui <input checked="" type="checkbox"/> Non Dans ce cas fournir une désignation d'inventeur(s) séparée		
8 RAPPORT DE RECHERCHE		Uniquement pour une demande de brevet (y compris division et transformation)		
Établissement immédiat ou établissement différé		<input checked="" type="checkbox"/> <input type="checkbox"/>		
Paiement échelonné de la redevance		Paiement en trois versements, uniquement pour les personnes physiques <input type="checkbox"/> Oui <input checked="" type="checkbox"/> Non		
9 RÉDUCTION DU TAUX DES REDEVANCES		Uniquement pour les personnes physiques <input type="checkbox"/> Requête pour la première fois pour cette invention (joindre un avis de non-imposition) <input type="checkbox"/> Requête antérieurement à ce dépôt (joindre une copie de la décision d'admission pour cette invention ou indiquer sa référence) :		
Si vous avez utilisé l'imprimé «Suite», indiquez le nombre de pages jointes				
10 SIGNATURE DU-DEMANDEUR OU DU MANDATAIRE (Nom et qualité du signataire)		VISA DE LA PRÉFECTURE OU DE L'INPI		
Pascale BROCHARD		 		

Procédé de reconnaissance vocale avec correction automatique

La présente invention concerne un procédé de reconnaissance vocale avec correction automatique dans les systèmes de reconnaissance vocale à syntaxe contrainte, c'est-à-dire que les phrases reconnaissables se trouvent dans un ensemble de possibilités déterminées. Ce procédé est particulièrement adapté à la reconnaissance vocale en milieu bruité, par exemple dans les cockpits d'avions d'arme ou civil, dans les hélicoptères ou dans l'automobile.

De nombreux travaux dans le domaine de la reconnaissance vocale à syntaxe contrainte ont permis d'obtenir des taux de reconnaissance de l'ordre de 95%, et ce, même dans l'environnement bruité d'un cockpit d'avion d'arme (environ 100-110 dBA autour du casque du pilote). Cependant, cette performance n'est pas suffisante pour faire de la commande vocale un média de commande primaire pour des paramètres critiques du point de vue de la sécurité de vol.

Une stratégie utilisée consiste à soumettre les commandes critiques à une validation du pilote, qui vérifie par la phrase reconnue, que les bonnes valeurs vont être affectées aux bons paramètres (« feedback primaire »). En cas d'erreur du système de reconnaissance – ou erreur de prononciation du pilote – le pilote doit énoncer à nouveau toute la phrase, et la probabilité d'erreur sur la reconnaissance de la phrase à nouveau prononcée est la même. Ainsi par exemple, si le pilote énonce « Select altitude two five five zero feet », le système effectue les algorithmes de reconnaissance et donne un retour visuel au pilote. En envisageant le cas où une erreur se produit, le système va par exemple proposer « SEL ALT 2 5 9 0 FT ». Dans un système classique, le pilote doit alors prononcer de nouveau toute la phrase, avec les mêmes probabilités d'erreur.

Un système de correction d'erreur meilleur en terme de taux de reconnaissance consiste à faire prononcer au pilote une phrase de correction qui sera reconnue comme telle. Par exemple, si l'on reprend l'exemple précédent, le pilote pourra prononcer « Correction third digit five ». Cependant cette méthode augmente la charge de travail du pilote dans le procédé de reconnaissance, ce qui n'est pas souhaitable.

L'invention propose un procédé de reconnaissance vocale qui met en œuvre une correction automatique de la phrase prononcée permettant

d'obtenir un taux de reconnaissance proche de 100%, sans augmentation de la charge du pilote.

- Pour cela, l'invention concerne un procédé de reconnaissance vocale d'un signal de parole prononcé par un locuteur avec correction automatique, comprenant notamment une étape de traitement dudit signal de parole délivrant un signal sous une forme compressée, une étape de reconnaissance de formes pour rechercher, à partir d'une syntaxe préenregistrée lors d'une phase préalable, une phrase de ladite syntaxe la plus proche dudit signal sous sa forme compressée, et caractérisé en ce qu'il
- comprend
- la mémorisation du signal sous sa forme compressée,
 - la génération d'une nouvelle syntaxe dans laquelle ladite phrase déterminée lors de l'étape de reconnaissance antérieure est supprimée,
 - la réitération de l'étape de reconnaissance de formes pour rechercher, à partir de la nouvelle syntaxe, une autre phrase la plus proche dudit signal mémorisé.

D'autres avantages et caractéristiques apparaîtront plus clairement à la lecture de la description qui suit, illustrée par les figures annexées qui représentent :

- la figure 1, le schéma de principe d'un système de reconnaissance vocale de type connu;
- la figure 2, le schéma d'un système de reconnaissance vocale du type de celui de la figure 1 mettant en œuvre le procédé selon l'invention ;
- la figure 3, un schéma illustrant la modification de la syntaxe dans le procédé selon l'invention.

Sur ces figures, les éléments identiques sont référencés par les mêmes repères.

La figure 1 présente le schéma de principe d'un système de reconnaissance vocale à syntaxe contrainte de type connu, par exemple un système embarqué dans un environnement fortement bruité. Dans un système à syntaxe contrainte mono locuteur, une phase d'apprentissage hors temps réel permet à un locuteur donné d'enregistrer un ensemble de

références acoustiques (mots) stockés dans un espace de références 10. La syntaxe 11 est formée d'un ensemble de phrases qui représentent l'ensemble des chemins ou transitions possibles entre les différents mots. Typiquement, quelques 300 mots sont enregistrés dans l'espace de
 5 référence qui forment typiquement 400 000 phrases possibles de la syntaxe.

Classiquement, un système de reconnaissance vocale comporte au moins trois blocs comme illustré sur la figure 1. Il comporte un bloc 12 d'acquisition du signal de parole (ou prise de son), un bloc 13 de traitement du signal et un bloc 14 de reconnaissance de formes. Une description
 10 détaillée de l'ensemble de ces blocs selon un mode de réalisation se trouve par exemple dans la demande de brevet français FR 2 808 917 au nom de la déposante.

De façon connue, le signal acoustique traité par le bloc de prise de son 12 est un signal de parole capté par un transducteur électroacoustique.
 15 Ce signal est numérisé par échantillonnage et découpé en un certain nombre de trames recouvrantes ou non, de même durée ou non. Dans le bloc 13 de traitement du signal, on associe classiquement chaque trame à un vecteur de paramètres qui traduit l'information acoustique contenue dans la trame. Il y a plusieurs méthodes pour déterminer un vecteur de paramètres. Un
 20 exemple classique de méthode est celle qui utilise les coefficients cepstraux de type MFCC (abréviation de l'expression anglo-saxonne « Mel Frequency Cepstral Coefficient »). Le bloc 13 permet de déterminer dans un premier temps l'énergie spectrale de chaque trame dans un certain nombre de canaux fréquentiels ou fenêtres. Il délivre pour chacune des trames une
 25 valeur d'énergie spectrale ou coefficient spectral par canal fréquentiel. Il effectue ensuite une compression des coefficients spectraux obtenus pour tenir compte du comportement du système auditif humain. Il effectue enfin une transformation des coefficients spectraux compressés, ces coefficients spectraux compressés transformés sont les paramètres du vecteur de
 30 paramètres recherché.

Le bloc 14 de reconnaissance de formes est relié à l'espace de références 10. Il compare la série des vecteurs de paramètres issue du bloc de traitement du signal aux références obtenues lors de la phase d'apprentissage, ces références traduisant les empreintes acoustiques de
 35 chaque mot, chaque phonème, plus généralement de chaque commande et

que l'on appellera de façon générique « phrase » dans la suite de la description. Puisque la reconnaissance de formes s'effectue par comparaison entre vecteurs de paramètres, on doit avoir à disposition ces vecteurs de paramètres de base. On les obtient de la même manière que
 5 pour les trames de signal utile, en calculant pour chaque trame de base son énergie spectrale dans un certain nombre de canaux fréquentiels et en utilisant des fenêtres de pondération identiques.

A l'issue de la dernière trame, ce qui correspond généralement à la fin d'une commande, la comparaison donne soit une distance entre la
 10 commande testée et des commandes de référence, la commande de référence présentant la distance la plus faible est reconnue, soit une probabilité pour que la série des vecteurs de paramètres appartiennent à une suite de phonèmes. Les algorithmes classiquement utilisés pendant la phase de reconnaissance de formes sont dans le premier cas de type DTW
 15 (abréviation de l'expression anglo-saxonne pour Dynamic Time Warping) ou, dans le second cas de type HMM (abréviation de l'expression anglo-saxonne Hidden Markov Models). Dans le cas d'un algorithme de type HMM, les références sont des fonctions gaussiennes associées chacune à un phonème et non à des séries de vecteurs de paramètres. Ces fonctions
 20 gaussiennes sont caractérisées par leur centre et leur écart-type. Ce centre et cet écart type dépendent des paramètres de toutes les trames du phonème, c'est à dire des coefficients spectraux compressés de toutes les trames du phonème.

Les signaux numériques représentant une phase reconnue sont
 25 transmis à un dispositif 15 qui réalise le couplage avec l'environnement, par exemple par affichage de la phrase reconnue sur le viseur tête haute d'un cockpit d'avion.

Comme cela a été précédemment expliqué, pour les commandes critiques, le pilote peut avoir à sa disposition un bouton de validation
 30 permettant l'exécution de la commande. Dans le cas où la phrase reconnue serait erronée, il doit généralement répéter la phrase avec une probabilité identique d'erreur.

Le procédé selon l'invention permet une correction automatique de grande efficacité et simple à mettre en œuvre. Son implantation dans un

système de reconnaissance vocale du type de la figure 1 est schématisée sur la figure 2.

Selon l'invention, à l'issue de la phase de traitement du signal 13, on mémorise (étape 16) le signal de parole sous sa forme compressée (ensemble des vecteurs de paramètres également appelés « cepstres »). Dès qu'une phrase est reconnue, on génère une nouvelle syntaxe (étape 17) dans laquelle la phrase reconnue n'est plus un chemin possible de la syntaxe. On réitère alors la phase de reconnaissance de formes avec le signal mémorisé mais sur la nouvelle syntaxe. Préférentiellement, la reconnaissance de formes est réitérée de manière systématique pour préparer une autre solution possible. Si le pilote détecte une erreur dans la commande reconnue, il appuie par exemple sur un bouton spécifique de correction, ou exerce un appui court ou un double clic sur l'alternat de commande vocale et le système lui propose la nouvelle solution trouvée lors de la réitération de la reconnaissance de formes. On réitère les étapes précédentes pour générer de nouvelles syntaxes qui interdisent toutes les solutions précédemment trouvées. Quand le pilote voit la solution qui correspond réellement à la phrase énoncée, il valide par un moyen quelconque (bouton, voix, etc.).

Reprenons l'exemple cité précédemment en tirant bénéfice de l'invention. Le pilote énonce selon cet exemple « Select altitude two five five zero feet ». Le système effectue les algorithmes de reconnaissance et, par exemple à cause du bruit ambiant, reconnaît « Select altitude two five nine zero feet ». Un feedback visuel est donné au pilote : « SEL ALT 2 5 9 0 FT ». Alors que le locuteur est en train de lire la phrase reconnue, le système anticipe une éventuelle erreur en générant de façon automatique une nouvelle syntaxe dans laquelle la phrase reconnue est supprimée et en réitérant l'étape de reconnaissance de formes.

La figure 3 illustre par un schéma simple, dans le cas de l'exemple précédent, la modification de la syntaxe permettant avec un algorithme de reconnaissance de formes de type DTW la recherche d'une nouvelle phrase. La phrase énoncée par le locuteur selon l'exemple précédente est « SEL ALT 2 5 5 0 FT ». Nous supposons que la phrase reconnue par la première phase de reconnaissance de formes est « SEL ALT 2 5 9 0 FT ». Cette première phase fait appeler à la syntaxe d'origine SYNT1, dans laquelle

toutes les combinaisons (ou chemins) sont possibles pour les quatre chiffres à reconnaître. Lors d'une deuxième phase de reconnaissance de formes, la phrase reconnue est écartée des combinaisons possibles, modifiant ainsi l'arbre syntaxique comme cela est illustré sur la figure 3. Une nouvelle

5 syntaxe est générée qui interdit le chemin correspondant à la solution reconnue. Une deuxième phase est alors reconnue. La phase de reconnaissance de formes peut être réitérée avec, à chaque fois, génération d'une nouvelle syntaxe qui reprend la syntaxe précédente mais dans laquelle est supprimée la phrase précédemment trouvée.

10 Dans un mode de fonctionnement possible, le pilote indique au système qu'il désire une correction (par exemple par un appui court de l'alternat commande vocale) et dès qu'une nouvelle solution est disponible, elle est affichée. La recherche automatique d'une nouvelle phrase s'arrête

15 par exemple lorsqu'une phrase reconnue est validée par le pilote. Dans notre exemple, il est probable que dès la deuxième phase de reconnaissance de formes, le pilote voit « SEL ALT 2 5 5 0 FT ». Il peut alors valider la commande. Dans la mesure où de nombreuses erreurs de reconnaissance sont dues à des confusions entre des mots proches (par exemple, five-nine), l'invention permet de corriger presque à coup sûr ces erreurs avec un

20 minimum de charge de travail supplémentaire pour le pilote et de façon très rapide du fait de l'anticipation sur la correction que peut effectuer le procédé selon l'invention.

En outre, en générant une nouvelle syntaxe et en réitérant l'étape de reconnaissance de formes sur la nouvelle syntaxe, on n'accroît pas la

25 complexité de l'arbre syntaxique. L'algorithme de traitement peut donc effectuer la reconnaissance avec un délai similaire à chaque itération, ce délai étant imperceptible pour le pilote du fait de l'anticipation de la correction.

REVENDECATIONS

1- Procédé de reconnaissance vocale d'un signal de parole prononcé par un locuteur avec correction automatique, comprenant
5 notamment une étape (13) de traitement dudit signal de parole délivrant un signal sous une forme compressée, une étape (14) de reconnaissance de formes pour rechercher, à partir d'une syntaxe (SYNT1) préenregistrée lors d'une phase préalable, une phrase de ladite syntaxe la plus proche dudit signal sous sa forme compressée, et caractérisé en ce qu'il comprend

- 10 - la mémorisation (16) du signal sous sa forme compressée,
- la génération (17) d'une nouvelle syntaxe (SYNT2) dans laquelle ladite phrase déterminée lors de l'étape de reconnaissance antérieure est supprimée,
- 15 - la réitération de l'étape de reconnaissance de formes pour rechercher, à partir de la nouvelle syntaxe, une autre phrase la plus proche dudit signal mémorisé.

2- Procédé de reconnaissance vocale selon la revendication 1, caractérisé en ce que la recherche d'une nouvelle phrase est réitérée de façon systématique pour anticiper la correction.

20 3- Procédé de reconnaissance vocale selon la revendication 2, caractérisé en ce que chaque nouvelle phrase reconnue est proposée au locuteur sur sa demande.

4- Procédé de reconnaissance vocale selon l'une des revendications 2 ou 3, caractérisé en ce que la recherche d'une nouvelle
25 phrase est stoppée par validation d'une phrase reconnue par le locuteur.

5- Procédé de reconnaissance vocale selon l'une des revendications précédentes, caractérisé en ce que l'étape (13) de traitement comprend :

- 30 - une étape de numérisation et de découpage en une suite de trames temporelles dudit signal acoustique,
- une phase de paramétrisation de trames temporelles contenant de la parole de manière à obtenir, par trame, un vecteur de paramètres dans le domaine fréquentiel, l'ensemble de ces vecteurs de paramètres formant ledit signal sous sa
35 forme compressée.

6- Procédé de reconnaissance vocale selon la revendication 5, caractérisé en ce que la reconnaissance de forme fait appel à un algorithme de type DTW.

5 7- Procédé de reconnaissance vocale selon la revendication 5, caractérisé en ce que la reconnaissance de forme fait appel à un algorithme de type HMM.

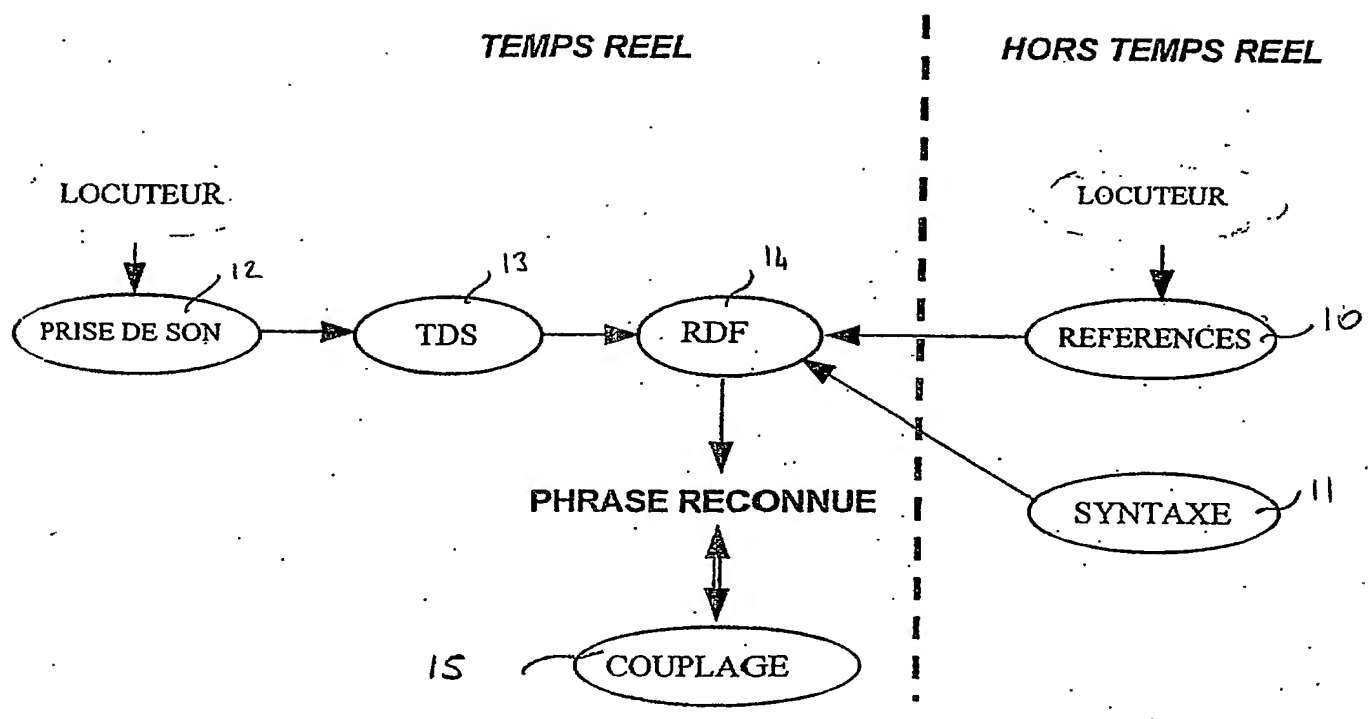


Fig. 1

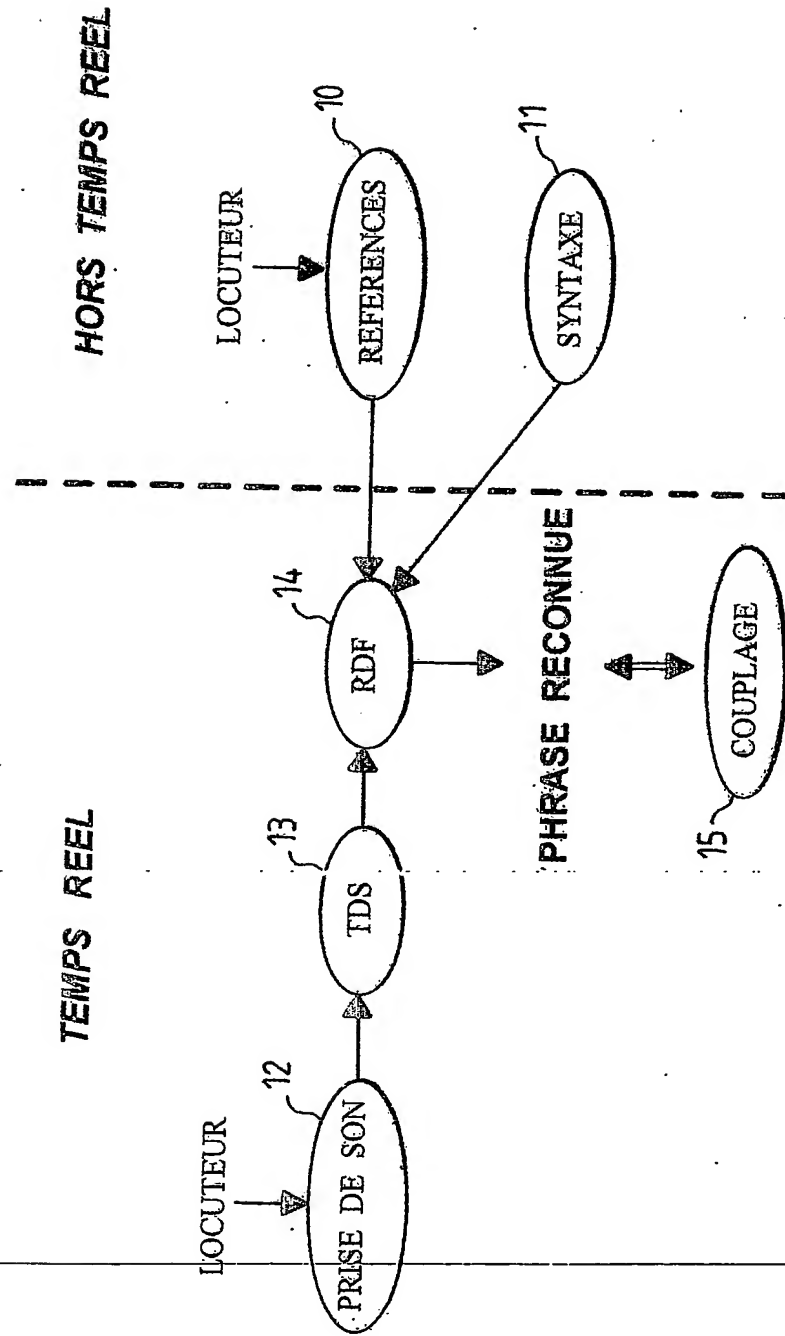


FIG.1

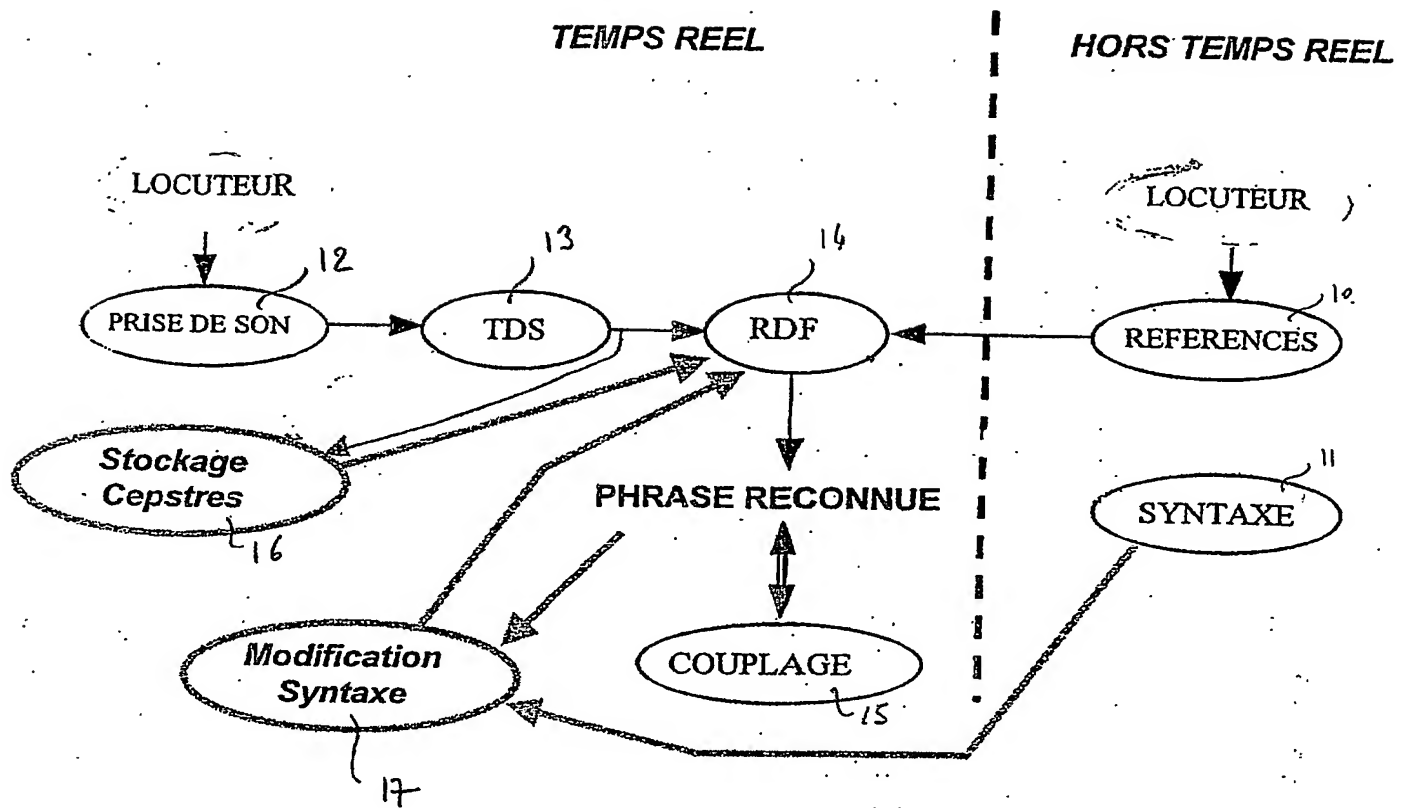


Fig. 2

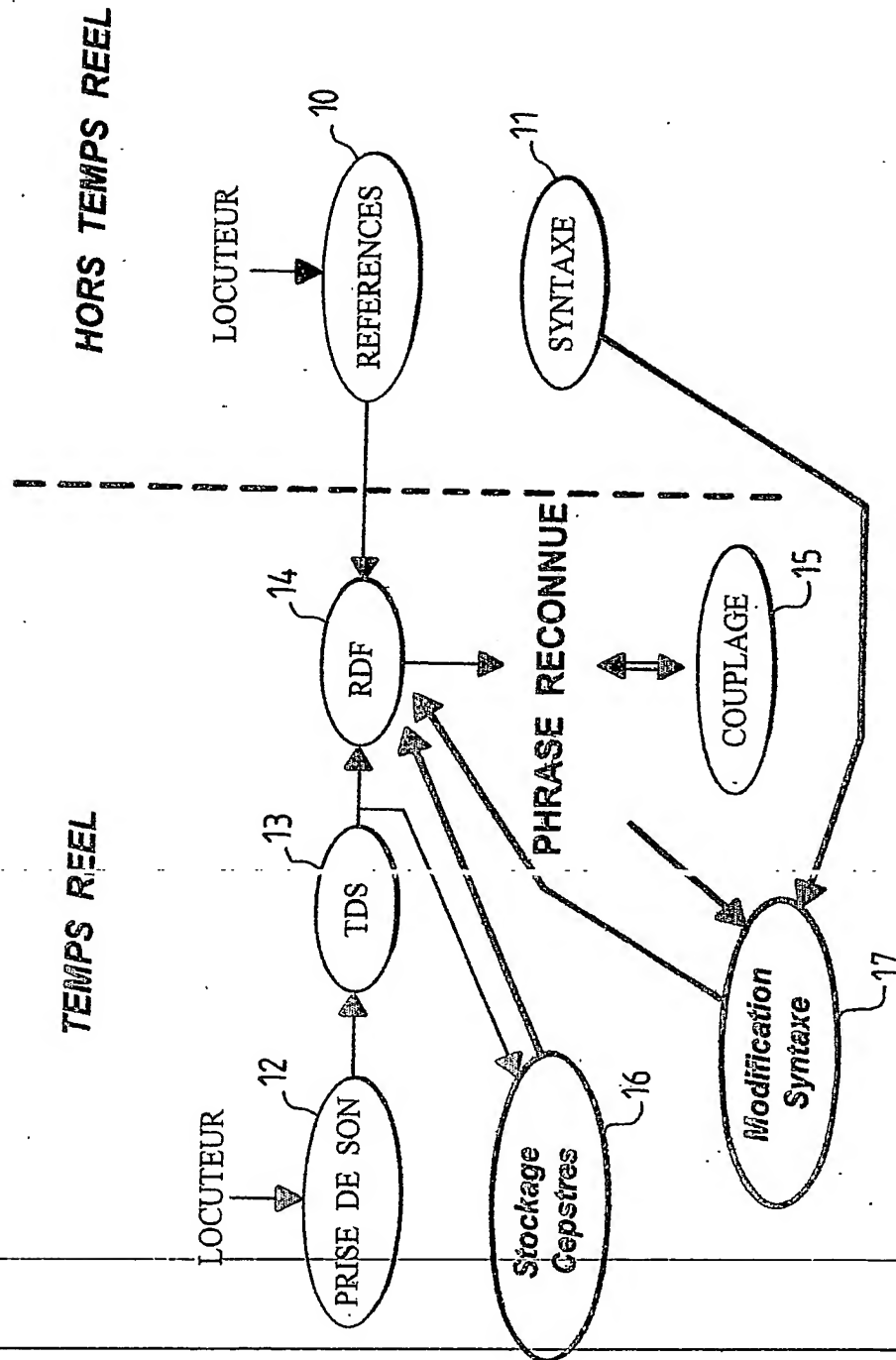
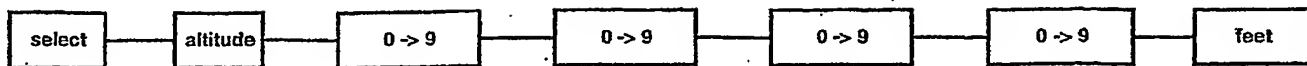


FIG.2

SYNT 1



SYNT 2

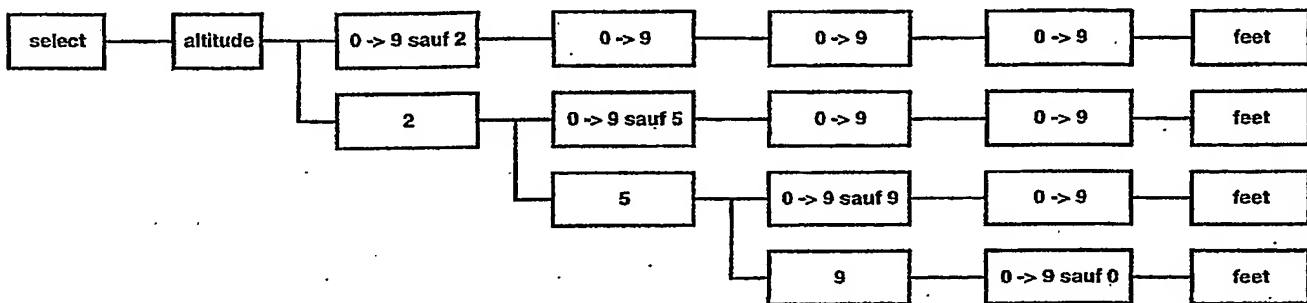


Fig. 3

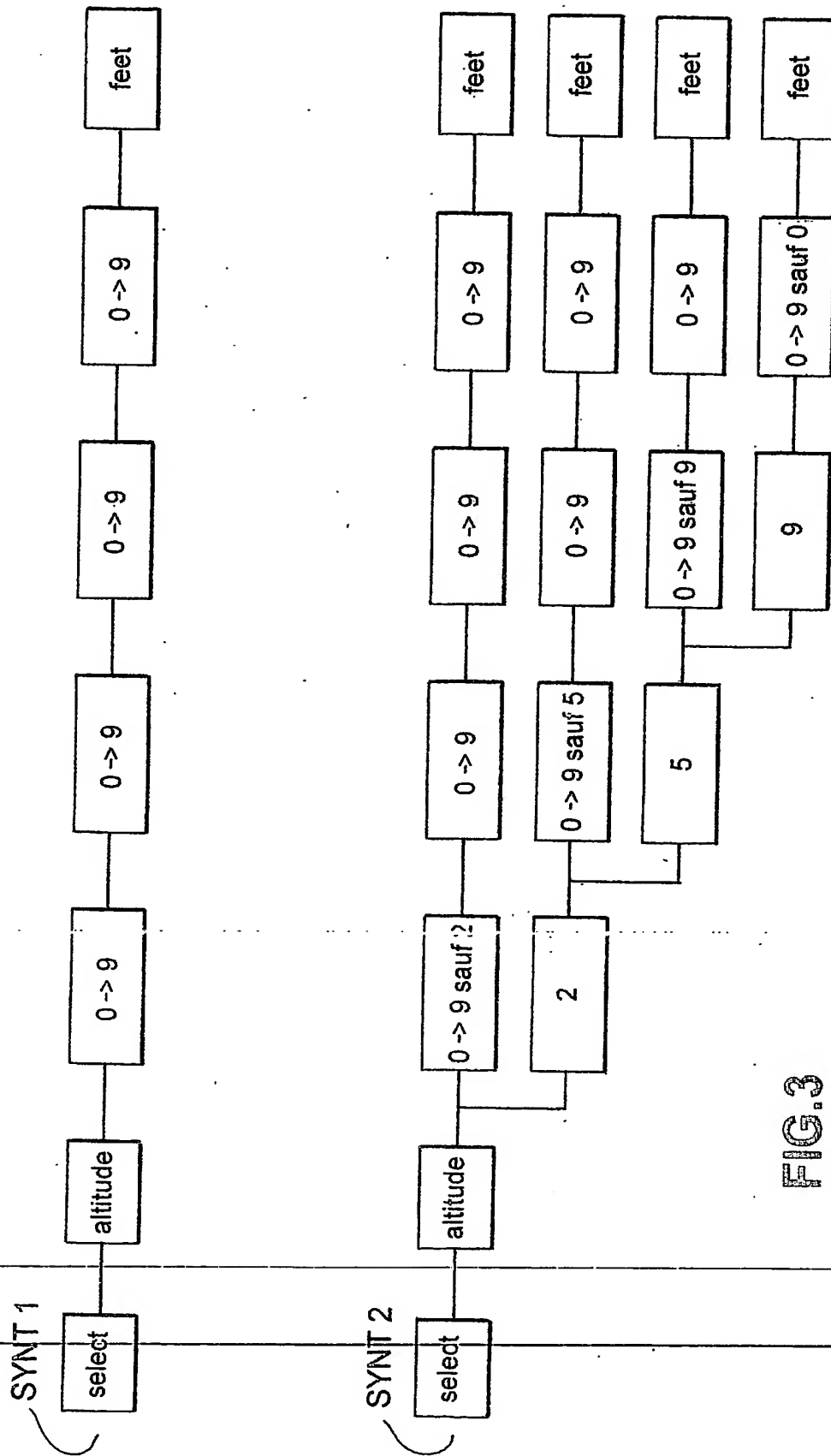


FIG. 3

ÉPARTEMENT DES BREVETS

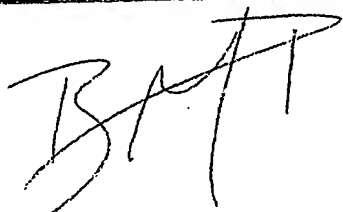
6 bis, rue de Saint Pétersbourg
5800 Paris Cedex 08

éléphone : 01 53 04 53 04 Télécopie : 01 42 93 59 30

DÉSIGNATION D'INVENTEUR(S) Page N° 1.. / 1..
(Si le demandeur n'est pas l'inventeur ou l'unique inventeur)

Cet imprimé est à remplir lisiblement à l'encre noire

DB 113 W / 260899

Vos références pour ce dossier (facultatif)		62877	
N° D'ENREGISTREMENT NATIONAL		02 M 789	
TITRE DE L'INVENTION (200 caractères ou espaces maximum)			
PROCÉDE DE RECONNAISSANCE VOCALE AVEC CORRECTION AUTOMATIQUE			
LE(S) DEMANDEUR(S) :			
THALES			
DESIGNE(NT) EN TANT QU'INVENTEUR(S) : (Indiquez en haut à droite «Page N° 1/1» S'il y a plus de trois inventeurs, utilisez un formulaire identique et numérotez chaque page en indiquant le nombre total de pages).			
Nom		POUSSIN	
Prénoms		Gilles	
Adresse	Rue	THALES Intellectual Property 13, avenue du Président Salvador Allende	
	Code postal et ville	94117	ARCUEIL Cedex
Société d'appartenance (facultatif)			
Nom			
Prénoms			
Adresse	Rue		
	Code postal et ville		
Société d'appartenance (facultatif)			
Nom			
Prénoms			
Adresse	Rue		
	Code postal et ville		
Société d'appartenance (facultatif)			
Nom			
Prénoms			
Adresse	Rue		
	Code postal et ville		
Société d'appartenance (facultatif)			
DATE ET SIGNATURE(S) DU (DES) DEMANDEUR(S) OU DU MANDATAIRE (Nom et qualité du signataire)			
Pascale BROCHARD 24 SEP. 2002			

PCT Application

FR0302770

